

SEQUENCE LISTING - WEE virus strain 71V-1658 - SEQ ID NO 1

LOCUS AF214040 11484 bp RNA VRL 27-JAN-2000
 DEFINITION Western equine encephalomyelitis virus strain 71V-1658
 nonstructural polyprotein and structural polyprotein genes,
 complete cds.
 ACCESSION AF214040 AF143811
 VERSION AF214040.1 GI:6760410
 KEYWORDS .
 SOURCE Western equine encephalomyelitis virus.
 ORGANISM Western equine encephalomyelitis virus Viruses; ssRNA
 positive-strand viruses, no DNA stage; Togaviridae; Alphavirus.
 REFERENCE 1 (bases 1 to 11484)
 AUTHORS Netolitzky,D.J., Schmaltz,F.L., Parker,M.D., Rayner,G.A.,
 Fisher,G.R., Trent,D.W., Bader,D.E. and Nagata,L.P.
 TITLE Complete genomic RNA sequence of western equine
 encephalitis virus and expression of the structural genes
 JOURNAL J. Gen. Virol. 81 Pt 1, 151-159 (2000)
 MEDLINE 20109043
 REFERENCE 2 (bases 1 to 11484)
 AUTHORS Netolitzky,D.J., Schmaltz,F.L., Parker,M.D., Rayner,G.A.,
 Fisher,G.R., Trent,D.W., Bader,D.E. and Nagata,L.P.
 TITLE Direct Submission
 JOURNAL Submitted (08-DEC-1999) Medical Countermeasures Section,
 Defence Research Establishment Suffield, P.O. Box 4000, Stn Main,
 Medicine Hat, Alberta T1A 8K6, Canada
 FEATURES Location/Qualifiers
 source 1..11484
 /organism="Western equine encephalomyelitis virus"
 /strain="71V-1658"
 /db_xref="taxon:11039"
 5'UTR <1..24
 misc_feature 1..11484
 /note="sequence of this entry added to the sequence of
 AF214039, strain CBA87, constitutes the Western equine
 encephalomyelitis virus complete genome"

 CDS 25..7428
 /note="encodes nsP1, nsP2, nsP3 and nsP4 proteins"
 /codon_start=1
 /transl_except=(pos:5581..5583,aa:OTHER)
 /product="nonstructural polyprotein"
 /protein_id="AAF28339.1"
 /db_xref="GI:6760411"

 /translation="MERIHVDLDADSPYVKSLQRTFPQFEIEARQVTDNDHANARAFS
 HVATKLIIESEVDRDQVILDIGSAPVRHAHSNHRYHCICPMISAEDPDRLQRYAERLKK

Feature "6482200"

NEQICLYILGESMSSIRSKCPVEESEASAPPHTLPCLCNYAMTAERVYRLRSAKKEQF
AVCSSFLLPKYRITGVQKLQCSKPVLFSGVVPPAVHPRKYAEIILETPPPPATTTVIC
EPTVPERIPSPVISRAPSAESLLSLGGVSFSSSATRSSTAWSYDRRFVVTADVHQAN
TSTWSIPSAPGLDVQLPSDVTDSHWSIPSASGFVVRTPSVQDLTAECAKPRGLAEIMQ
DFNTAPFQFLSDYRPVPAPRRRPIPSPRSTASAPPVVKPRRTKYQQPPGVARAISEAE
LDEYIRQHSNXRYEAGAYIFSSETGQGHLLQOKSVRQCKLQEPILDRAVHEKYYAPRLD
LEREKMLQKKLQLCASEGNRSRYQSRKVENMKAITAERLISGLGTYLSSEVNPVECYR
VNYPVPIYSSTVINRFTSAEVAVKTCNLVIQENYPTVASYCITDEYDAYLDMVDGASC
CLDTATFCPAKLRSYPKKHLSYLQPEIRSAVPSPIQNTLQNVLAATAKRN CNVTQMREL
PVLDSAAFNVDCFKKYACNDEYWDTFRDNP IRLTTENV TQYVTKLKGPKAAALFANTH
NLKPLQEI PMDQFVMDLKR DVKVTPGT KHT EERPKVQVIQAADPLATAYLCGIHREL V
RRLNAVLLPNIHTLFDMSAEDFDAI IAEHFHHGDPVLETDIASFDKSEDDAIAISALM
ILEDLGVDQPLLDLIEAAFGNITSVHLPTGTRFKFGAMMKSGMFLTLFVNTLVNIMIA
SRVLRERLTTSACAASIGDDNIVHGVVSDTLMAERCATWLNMEVKIIDAVIGIKAPYF
CGGFILVDQITGTACRVADPLKRLFKLGKPLPVDDTQDCDRRRALHDEAMRWNRRIGIT
DELVKAVESRYEIIILAGLIITSLSTLAESVKNFKSIRGSPITLYG"

CDS 7473..11183
/note="encodes nucleocapsid, E3, E2, 6K and E1 proteins"
/codon_start=1
/product="structural polyprotein"
/protein_id="AAF28340.1"
/db_xref="GI:6760412"

/translation="MFPYPQLNFPPVYPTNPMAYRDPNPPRRRWRPFRPPLAAQIEDL
RRSIVNLTFKQRSPNPPPGPPPKKKKSAPKPKPTQPKKKKQQA KRTKRKPKPKGRQRM
CMKLES DKTFPIMLNGQVNGYACVVGGRLMKPLHVEGKIDNEQLAAVKLKKASMYDLE
YGDVPQNMKSDTLQYTS DKPPGFYNWHGAVQYENGRFTVPRGVGGKGDSGRPI LDNR
GRVVAIVLGGANEGTRTALS VVTWNQKGV TIRDTPEGSEPWSLVTALCVLSNVTFPCD

KPPVCYSLTPERTLDVLEENVDPNPYDTLLENVLKCPSSRRPKRSITDDFTLTSPYLGF
 CPYCRHSTPCFSPIKIENVWDESDDGSIRIQVSAQFGYNQAGTADVTKFRYMSFDHHDH
 DIKEDSMEKIAISTSGPCRRLGHKG YFLLAQCPPGDSVTVSITSGASENSCTVEKKIR
 RKFVGREEYLFPPVHGKLVKCHVYDHLKETSAGYITMHRPGPHAYKSYLEEASGEVYI
 KPPSGKNVTYECKCGDYSTGIVSTRTKMNGCTKAKQCIAYKSDQTKWVFNSPDLIRHT
 DHSVQGLHIPFRLTPTVCPVPLAHTPTVTKWFKGITLHLTAMRPTLLTTRKLGLRAD
 ATAEWITGSTSRNFSVGREGLEYYVWGNHEPVRVWAQESAPGDPHGWPHEII IHYYHRH
 PVYTVIVLCGVALAILVGTASSAACIAKARRDCLTPYALAPNATVPTALAVLCCIRPT
 NAETFGETLNHLWFNNQPFLWAQLCIPLAALVILFRCFSCCMPFLLVAGVCLGKVDAF
 EHATTVPNVPGIPYKALVERAGYAPLNLEITVVSSSELTPTSTNKEYVTCKFHTVIPSPQ
 VKCCGSLECKASSKADYTCRVFGGVYPFMWGAQCFCDSSENTQLSEAYVEFAPDCTID
 HAVALKVHTAALKVGLRIVYGNTTAHLDTFVNGVTPGSSRDLKVIAGPISAAFSFPDH
 KVVIRKGLVYNYDFPEYGAMKPGAFGDIQASSLDATDIVARTDIRLLKPSVKNIHVPY
 TQAVSGYEMWKNNNGRPLQETAPFGCKIEVEPLRASNCAYGHIPI SIDIPDAAFVRSS
 ESPTILEVSVCTVADCIYSADFGGSLTLQYKADREGHCPVHSHSTTAVLKEATHTHTAV
 GSITLHFSTSSPQANFIVSLCGKKSTCNAECKPPADHII GEPHKVDQEFQAAVSKTSW
 NWLLALFGGASSLIVVGLIVLVCCSSMLINTRR"

3'UTR 11184..11484

BASE COUNT 3251 a 2842 c 2808 g 2583 t
 ORIGIN

```

1 accctacaaa ctaatcgatc caatatggaa agaattcacg ttgacttaga tgctgacagc
61 ccgtatgtca agtcggtaca gcggacgttt ccacaatttg agatcgaagc aaggcaggtc
121 actgacaatg accatgccaa tgccagagcg ttttcgcatg tggcaacaaa gctcattgag
181 agcgaagtcg accgggacca agttatcttg gacattggaa gtgcgcccgt cagacatgca
241 cattccaatc accgctatca ttgtatctgc cctatgataa gcgctgaaga cccggacaga
301 ctacaacggt atgcagaaag acttaagaaa agtgacatta ccgacaagaa catagcctct
361 aaggcggcag acctgctgga agtcatgtca acaccagacg cagagactcc atctctgtgt
421 atgcacacag acgccacgtg taggtacttt ggaagtgtag cagtatacca agatgtgtac
481 gcagtccatg caccgacatc aatctaccac caggcgctta aaggagttag gacaatttac
541 tggataggct ttgacacgac cccttttatg tacaaaaaca tggcagggtc ctaccctact

```

601	tacaacacga	actgggctga	cgagagagta	ttggaagcac	gtaacattgg	cctcggtaac
661	tcagatcttc	aggagagcag	gcttggaaaa	ctctcaatcc	ttaggaagaa	gaggctccaa
721	cctactaata	agatcatatt	ctcggttggg	tcaacaatct	acacagaaga	tagatcactg
781	ttacgtagct	ggcatcttcc	aaacgtgttc	cacttgaaag	gaaagtctaa	cttcacaggt
841	agatgtggga	ccattgtcag	ctgtgaaggg	tacgtcatca	aaaagataac	gatcagccca
901	ggactatacg	gtaaagttga	gaacttggcg	tccacaatgc	atcgcgaggg	tttcttgagt
961	tgcaaagtca	cagatacgct	gcgcggcgag	agggtttctt	ttgctgtgtg	tacgtatgta
1021	ccagccacac	tttgcgatca	gatgacaggg	attctggcaa	ctgacgttag	tgtggatgac
1081	gcacaaaaac	tattggttgg	gctcaaccaa	aggattgtcg	tcaatggtag	gacgcaaaga
1141	aatactaaca	caatgcagaa	ctatctatta	ccagtggtcg	cccaggcggt	ttccaggtgg
1201	gcgcgtgaac	atcgtgccga	cttggacgac	gagaaagaac	taggggtgcg	ggagcgcact
1261	cttactatgg	gctgctgctg	ggctttcaag	accagaaaa	tcacatccat	ctacaagaag
1321	cctggtagcg	aaacaattaa	gaaagtacct	gccgtctttg	actcatttgt	gattccacgc
1381	cttaccagcc	acgggctcga	tatgggcttc	cgccgtaggg	tcaagctgct	gcttgaacca
1441	actgtcaaac	ccgcaccggc	tattacaatg	gccgatgtgg	agcatctgcg	tggcttacag
1501	caagaagctg	aagaagtggc	tgcagcggaa	gagatcagag	aagccctgcc	acccttgctc
1561	cctgaaatag	aaaaagagac	cgtagaggca	gaagtagacc	tcattatgca	agaggcagga
1621	gcaggtagcg	tggagacacc	acgaggacac	atcaggggtga	caagttaccc	aggcgaagag
1681	aagattgggt	cttacgctat	actttcaccc	caggcggtat	tgaatagtga	aaaactggcg
1741	tgtatccacc	cattggcgga	acaagtactg	gtaatgactc	acaaaggtag	ggcagggaga
1801	tacaaagtcg	agccatacca	cggtaaggtc	attgtaccag	aagggacggc	ggtccctggt
1861	caagacttcc	aggcattgag	tgagagcgct	acgatcgttt	tcaacgagag	ggagtccgta
1921	aacagatacc	tgcaccacat	cgcaatcaac	ggaggagcgc	taaacactga	cgaagagtac
1981	tataagactg	taaagactca	ggacacagac	tcagaatacg	tcttcgatat	tgacgcacga
2041	aagtgtgtta	agcgagaaga	cgcagggtccc	ttgtgcctaa	ccggtgatct	ggtagatcca
2101	ccatttccacg	agtttgcgta	cgagagtctc	aagacacgac	cagcagcacc	tcacaaagtc
2161	ccaaccatcg	gagtctatgg	agtgccagggt	tcaggtaaata	ctggaatcat	caaaagcgct
2221	gtgactaaga	aagatctggg	tgtgagtgcg	aagaaggaaa	actgcgcgaga	aatcatcagg
2281	gatgtaagga	ggatgagacg	tatggatggt	gctgctagga	ctgtcgattc	agtgtctcta
2341	aatgggggtta	agcaccctcg	taacactctg	tacattgatg	aggcatttgc	ctgccatgca
2401	gggacgctgc	tggcactgat	tgccatcgtc	aaacctaaga	aagtgggtatt	gtgcgggggac
2461	ccaaaacaat	gcggcttctt	taacatgatg	tgccgtgaaag	tacatttttaa	ccatgacata
2521	tgcactgaag	tgtaccataa	aagcatctct	aggagggtgca	cacagactgt	aaccgccatc
2581	gtctccacgc	tcttctacga	caagcgaatg	aagacgggtta	acctatgtgc	tgataaaaac
2641	atcatagata	ccacagggac	cacaaagccg	cacaaagatg	atctgattct	aacctgtttc
2701	agaggatggg	tgaacacgct	acagattgac	tacaaaaatc	acgaaatcat	gactgctggc
2761	gcatcgcaag	gacttacgcg	gaaaggcggt	tatgctgtca	ggtacaaaagt	caacgagaat
2821	ccactctact	cgcagacttc	tgagcacgtg	aacgtgttac	ttacacgcac	agaaaaacgc
2881	attgtctgga	agacgctagc	tggatgaccc	tggataaaga	cacttacagc	taaatatccc
2941	ggggattttca	cggcttcatt	ggacgactgg	cagcgcgaac	acgacgccat	tatggcacgc
3001	gttcttgata	agcgcgagac	agctgatgtg	ttccagaata	aggtgaacgt	ctgctgggcg
3061	aaggcttttag	agccagtctt	ggccacggcc	aacattgtgc	tgacgagaca	gcagtgggag
3121	acgttgcacc	cattcaagca	tgacagagcg	tactcacctg	aaatggcact	gaacttcttt
3181	tgcaccaggt	tctttggagt	agacctggac	agtgggttat	tttccgctcc	taccgtcgca
3241	cttacttaca	gggatcagca	ctgggataac	tcgccaggga	agaacatgta	tgggcttaat
3301	agagaggtag	caaaggagtt	gtcacggcga	tatccgtgca	tcacaaaagc	ggttgacaca
3361	ggcagggtag	ctgatataag	gaataatacc	atcaaggact	actctccaac	aattaatgtg
3421	gttccattaa	atcgccgggt	gccccactcg	ttgatcggtg	accacaaagg	acaggggtaca
3481	actgatcaca	gcggattcct	atctaagatg	aagggcaaat	ctgtgttggt	gatcggcgat
3541	cctatcagca	ttccagggaa	gaaagtagag	tccatgggtc	cattgcccac	taataaccatc

3601	aggtgtgatc	tcgatttggg	aatacctage	catgtcggta	aatatgacat	tatctttgtc
3661	aatgttagga	ccccgtacag	gaaccatcac	taccaacagt	gcgaggatca	cgctatccac
3721	cacagcatgc	taacgtgtaa	ggctgtccac	cacctgaaca	ctggcggaac	atgtgtggct
3781	atagggatatg	ggcttgetga	tcgcgcaacc	gagaatatca	tcactgcggt	ggcacgctca
3841	tttaggttta	cccgtgtctg	tcagcctaag	aacactgccg	aaaatactga	ggttctcttc
3901	gtgttcttcg	gcaaggacaa	cggcaaccac	acacatgacc	aggacagact	cgggtgtagt
3961	cttgacaaca	tctatcaagg	gtcaaccagg	tacgaggcag	ggagagctcc	agcgtacaga
4021	gtgatcagag	gtgacattag	caagagcgct	gaccaagcta	tcgttaatgc	tgctaatagc
4081	aaaggtcaac	cagggtcccg	agtgtgcggt	gcactgtacc	gaaaatggcc	ggctgctttt
4141	gatagacagc	caatagctgt	cgggacggct	agacttgtga	agcacgaacc	gctcatcata
4201	catgctgtag	gacccaattt	ttctaagatg	ccggaaccgg	agggcgacct	taagctcgca
4261	gctgcctaca	tgagcatagc	gtccatcgct	aacgctgagc	ggattacaaa	aatatcagta
4321	ccgctactgt	caaccggcat	ctattctggt	ggcaaagatc	gagtgatgca	atcattgcat
4381	cacctgttca	ctgcttttega	cactacggat	gccgatgtca	ccatatattg	cttgataaaa
4441	caatgggaga	ccaggataat	cgaggccatt	caccgcaaag	aaagcgtcga	aattctggat
4501	gatgacaagc	cagtagacat	tgacttggtc	aggggccacc	caaacagctc	tttggcaggc
4561	agaccagggt	actccgtcaa	tgagggcaag	ttgtattcat	acctggaagg	tacacgattc
4621	catcagaccg	ccaaggacat	tgccgaaatc	catgcaatgt	ggcccaacaa	atctgaggct
4681	aatgagcaga	tttgcttgta	catcctgggg	gagagtatgt	ccagcatccg	ctccaaatgc
4741	ccagtagagg	agtcagaggc	gtctgctcca	cctcacacac	ttccatgcct	gtgtaattac
4801	gctatgacgg	ctgagcgcgt	atacagggtg	cgctctgcga	agaaagaaca	gttcgccgta
4861	tgctcatcat	tctgttgcc	gaagtacagg	atcacaggcg	tgcalaagct	acagtgcagc
4921	aaaccagtc	tgttttcagg	cgctcgta	ccggctgtac	accccaggaa	gtacgcggaa
4981	ataattctag	aaacgccacc	accgccagca	acgacaaccg	taatattgtga	accaactgtg
5041	ccagaacgta	taccagtc	ggtgatttct	agagcaccaa	gtgcggaatc	actgctatcg
5101	cttggcgcg	tctcgttctc	tagctctgcc	acacgctcgt	caaccgcctg	gagcgactat
5161	gacaggcggt	ttgtggttac	agctgatgtg	catcaagcga	acacgtctac	gtggagcatc
5221	cctagtgtc	ctggcttgga	cgtccagctg	ccttctgacg	tcactgattc	ccactggagt
5281	attccaagt	catcaggctt	tgaagtgaga	acaccatctg	tacaggacct	aactgcggag
5341	tgtgcgaagc	ctcgtggact	ggccgaaata	atgcaagact	tcaatactgc	tcctttccag
5401	tttctttcgg	actacagacc	agtaccggca	ccacggagac	gccccatccc	atcacctaga
5461	tcgacggctt	ccgcacctcc	agttccaaag	ccacgcagga	ctaagtacca	acaaccacca
5521	ggagtcgcta	gagcgatctc	agaagcggag	ttggacgagt	acatccgtca	acactccaac
5581	tgacgggatg	aagcgggagc	gtatattttc	tcacgga	caggccaagg	tcacctcaa
5641	cagaaatcag	tacgtcaatg	taaactacaa	gaacctatat	tggatcgggc	cgtccatgag
5701	aagtattacg	ccccgcgcct	cgatctcgaa	agagagaaaa	tggttacagaa	gaaactgcaa
5761	ttatgcgcct	ctgaaggaaa	tagaagcagg	tatcaatcac	gaaaagtaga	aaatatgaaa
5821	gcaattacag	cggagcgact	catttctgga	ttgggcacat	atctatcatc	agaagtgaat
5881	cctgtcgagt	gttacagagt	caattatcct	gtaccaatct	actcgtcaac	ggtaattaac
5941	aggtttacat	ctgcagagg	cgcggttaaa	acgtgcaact	tagttatcca	agagaattac
6001	cctacagtag	ccagttattg	tataacagat	gaatacagatg	cgtatcttga	catggtggac
6061	ggcgcatcgt	gctgtctaga	tacagccact	ttttgtccgg	ctaaactgag	aagctaccca
6121	aagaagcata	gctatttgca	gccagagata	agatcagccg	tcccatcgcc	tatacagaat
6181	acattacaaa	atgtattggc	tgacgctact	aaaagggaatt	gcaacgttac	ccaaatgcga
6241	gaattacctg	tcttagattc	ggcggcattt	aatgttgatt	gtttcaagaa	atacgcattg
6301	aatgatgagt	actgggatac	ctttcgcgat	aaccctattc	ggctaactac	agagaacggt
6361	acgcaatatg	tgacaaagct	gaaagggccg	aaagcagcag	cattgtttgc	gaatactcat
6421	aatctaaaac	cgttgcagga	gataccaatg	gatcaattcg	tcattggatct	aaagagagat
6481	gtcaaagtta	ctcccgccac	gaaacataca	gaggagcggc	ctaagggtgca	ggttattcag
6541	gctgcagatc	cccttgctac	cgtttacctt	tgccgggatcc	atcgggaatt	agtcctgtaga

6601	ctgaatgcgg	tgetttctgcc	aaatatccat	actctcttcg	acatgtcagc	ggaagatttt
6661	gatgcgatta	ttgctgaaca	tttccaccac	ggcgacccag	tattggaaac	ggacatcgcg
6721	tcgtttgata	aaagcgaaga	cgacgctatc	gccatttcgg	cgttgatgat	ccttgaggac
6781	ttaggtgtcg	accaaccgct	cttagatttg	atagaggcgg	cgttcggcaa	tatcacatct
6841	gtgcacctac	ctacaggaac	gaggtttaaa	tttggtgcca	tgatgaaatc	cggtatgttc
6901	ttaacgctgt	ttgtcaacac	actagtcaat	atcatgattg	ctagcagagt	actacgtgaa
6961	cggttaacca	cgtcagcgtg	cgcggcctct	atcggcgacg	ataacatagt	gcatgggtgc
7021	gtctccgaca	ccttgatggc	ggagagatgc	gccacttggc	tgaacatgga	agtaaaaatt
7081	attgatgcag	ttattggtat	caaagcacc	tacttctgtg	ggggatttat	cctgggtggac
7141	cagataacag	gcacagcctg	cagagtcgca	gaccctctaa	aaaggctttt	taagcttgga
7201	aaaccattgc	cagtcgatga	tacccaagac	tgcgaccgcc	gccgggcact	gcatgatgaa
7261	gcaatgcgat	ggaacagaat	tgggaattacg	gacgagttag	tgaaggccgt	agaatccaga
7321	tacgagatca	tactggcagg	cctgatcatc	acgtctctgt	ccacgttagc	cgaaagcggt
7381	aagaacttca	agagcataag	agggagccca	atcaccctct	acggctgacc	taaatagggtg
7441	acgtagtaga	cacgcaccta	cccaccggca	gaatgtttcc	ataccctcag	ctgaactttc
7501	caccagttta	ccctacaaat	ccgatggcct	accgagatcc	aaaccctcct	aggcgccgct
7561	ggaggccggt	tcggccccc	ctggctgctc	aaatcgaaga	tcttaggagg	tcgatagtca
7621	acttgacttt	caaacaacga	tcacctaata	cgccgccagg	tccaccgcca	aagaagaaga
7681	agagtgtccc	taagccaaaa	cctactcagc	ctaaaaagaa	gaagcagcaa	gccaaagagga
7741	cgaaacgcaa	gcctaaacca	gggaaacgac	aacgtatgtg	tatgaagttg	gagtcggaca
7801	agacatttcc	gatcatgctg	aacggccaag	tgaatggata	tgcttgcgtt	gtcggaggaa
7861	ggctgatgaa	accactccac	gttgaaggaa	aaattgataa	tgagcaatta	gcggccgtga
7921	aattgaagaa	ggctagcatg	tacgacttgg	agtacggcga	cgttccccag	aacatgaaat
7981	cagacacgct	gcagtacacc	agcgacaaac	caccgggctt	ctacaactgg	caccacggcg
8041	cagtccagta	tgagaatggg	agatttaccg	taccgagagg	agtggggcgg	aaaggcgaca
8101	gcggaagacc	gatcctggac	aacagaggca	gagttgtggc	tattgttcta	ggaggtgcaa
8161	atgagggcac	gcgtacggcg	ctttcagtg	tcacttggaa	ccagaaagg	gtgaccatta
8221	gggatacccc	cgaaggttct	gaaccgtgg	cactagttac	agcgctatgc	gtgctttcga
8281	atgtcacgtt	cccattgcgac	aaaccacccg	tgtgctatc	actgacgcca	gaacgaacac
8341	tcgacgtgct	cgaagagaac	gtcgacaatc	caaattacga	cacgctgctg	gagaacgtct
8401	tgaaatgtcc	atcacgccgg	cccaaacgaa	gcattaccga	tgacttcaca	ctgaccagtc
8461	cctacctggg	gttctgccc	tattgcagac	actcaacgcc	gtgtttcagc	ccaataaaaa
8521	ttgagaacgt	gtgggacgaa	tctgatgatg	gatcgattag	aatccagggtc	tcgggcacaat
8581	tcggctacaa	tcaggcaggc	actgcggatg	tcaccaaatt	cgttacatg	tctttcgacc
8641	accaccatga	catcaaggaa	gacagtatgg	agaaaaatgc	tatcagcaca	tctggaccct
8701	gccgtcgtct	tggccacaaa	gggtacttcc	tgtagctca	atgtcctcca	ggtgacagt
8761	taaccgtcag	tatcacgagc	ggagcatctg	agaattcatg	caccgtggag	aaaaagatca
8821	ggaggaagtt	tgtcggtaga	gaggagtact	tgttcccacc	cgtccatgga	aagctggtaa
8881	agtgccacgt	ttacgatcac	ttgaaggaga	cgtctgccgg	gtacataacc	atgcacaggc
8941	caggcccaca	cgcgtataag	tcctatctgg	aggaagcgtc	aggcgaagt	tacattaaac
9001	caccttctgg	caagaacgtc	acctacgaat	gtaagtgtgg	cgactacagc	acagggtatcg
9061	tgagcacgcg	aacgaagatg	aacggctgca	ctaaagcaaa	acagtgcatt	gcctacaaga
9121	gcgaccaaac	gaaatgggtc	ttcaactcgc	cggatcttat	taggcacaca	gaccactcag
9181	tgcaaggtaa	attgcacatt	ccattccgct	tgacaccgac	agtctgccc	gttccggttag
9241	ctcacacgcc	tacagtacag	aagtggttca	aaggcatcac	cctccacctg	actgcaatgc
9301	gaccaacatt	gctgacaacg	agaaaattgg	ggctgcgagc	agacgcaaca	gcagaatgga
9361	ttacagggtc	tacatccagg	aatttttctg	tggggcgaga	agggtgagg	tacgtatggg
9421	gtaaccatga	accagtacga	gtctgggccc	aggagtccgc	accaggcgac	ccacatggat
9481	ggccgcgatga	gatcatcatc	cactattatc	atcggcaccc	agtctacact	gtcattgtgc
9541	tgtgtgggtg	cgctcttgct	atcctggtag	gcactgcac	atcagcagct	tgcategcca

9601	aagcaagaag	agactgcttg	acgccatacg	cgcttgccac	gaacgcaacg	gtacccacag
9661	cattagcggg	tttgtgctgc	attcggccaa	ccaacgctga	aacattttgga	gaaactttga
9721	accatctgtg	gtttaacaac	caaccgtttc	tctgggcaca	gttgtgcatt	cctctggcag
9781	cgcttgttat	tctgttccgc	tgtttttcat	gctgcatgcc	ttttttattg	gttgcaggcg
9841	tctgcctggg	gaaggtagac	gccttcgaac	atgcgaccac	tgtgccaaat	gttccgggga
9901	tcccgataaa	ggcggttggtc	gaacgcgcag	gttacgcgcc	acttaacctg	gagatcacgg
9961	tcgtctcatc	ggaattaaca	ccttcaacta	acaaggagta	cgtgacctgc	aaattccaca
10021	cagtcattcc	ttcaccacaa	gttaaattgct	gcgggtccct	cgagtgcgaag	gcattcctcaa
10081	aggcggatta	cacatgccgc	gtttttggcg	gtgtgtaccc	tttcatgtgg	ggaggcgcac
10141	aatgcttctg	tgacagtga	aacacacaac	tgagtgaggc	gtacgtcgag	ttcgctccag
10201	actgcactat	agatcacgca	gtcgacttaa	aagttcacac	agctgctctg	aaagtccggc
10261	tgcgtatagt	atacggcaac	accaccgcgc	acctggatac	gtttgtcaat	ggcgtcacgc
10321	caggttccctc	acgggacctg	aaggctcatag	cagggccgat	atcagccgct	ttttcaccct
10381	ttgaccataa	ggtcgtcatc	agaaaggggc	ttgtttacaa	ctacgacttc	cctgagtatg
10441	gagctatgaa	accaggagcg	ttcggcgata	ttcaagcatc	ctcgcttgat	gctacagaca
10501	tagtagcccg	cactgacata	cggctgctga	agccttctgt	caagaacatc	cacgtccctc
10561	acacccaagc	agtatcaggg	tatgaaatgt	ggaagaacaa	ctcaggacga	cccctgcaag
10621	aaacagcacc	atltggatgt	aaaattgaag	tggagcctct	gcgagcgtct	aactgtgctt
10681	acgggcacat	ccctatctcg	attgacatcc	ctgatgcagc	ttttgtgaga	tcatacagaat
10741	caccaacaat	tttagaagtt	agctgcacag	tagcagactg	catttattct	gcagactttg
10801	gtggttctct	aacattacag	tacaaagctg	acaggggagg	acattgtcca	gttactctcc
10861	actccacgac	agctgttttg	aagggaagcg	ccacacatgt	gactgccgta	ggcagcataa
10921	cactacattt	tagcacatcg	agcccacaag	caaattttat	agtttcgcta	tgccggcaaga
10981	agtccacctg	caatgctgaa	tgtaaaccac	cggccgacca	cataattgga	gaaccacata
11041	aagtcgacca	agaattccag	gcggcagttt	ccaaaacatc	ttggaactgg	ctgcttgcac
11101	tgtttggggg	agcatcatcc	ctcattgttg	taggacttat	agtgttggtc	tgcagctcta
11161	tgcttataaa	cacacgtaga	tgactgagcg	cggacactga	catagcggta	aaactcgatg
11221	tacttccgag	gaagcggtgt	gcataatgcc	acgcgccgct	tgacactaaa	actcgatgta
11281	tttccgagga	agcacagtgc	ataatgctgt	gcagtgtcac	attaatcgta	tatcacacta
11341	catattaaca	acactatatc	actttttatga	gactcactat	gggtctctaa	tatacactac
11401	acatatttta	cttaaaaaaca	ctatacacac	tttataaatt	ctttttataat	ttttcttttg
11461	tttttatttt	gttttttaaaa	tttc			

SEQUENCE LISTING - pcDWXH-7 - SEQ ID NO 2

<210> SEQ ID NO 2
 <211> LENGTH: 4150
 <212> TYPE: DNA
 <213> ORGANISM: WESTERN EQUINE ENCEPHALITIS VIRUS
 <214> FEATURES:

Location/Qualifiers

source 1..4150
 /organism="Western equine encephalomyelitis virus"
 /strain="71V-1658"
 vector sequence 1-9
 5' SacI primer 9-20
 3' end of NS4 gene 16-114
 intragenic region 115-159
 CDS 160-3869
 /note="encodes capsid, E3, E2, 6K and E1 proteins"
 /codon_start=160
 /product="structural polyprotein"
 /translation="MFPYPQLNFPPVYPTNPMAYRDPNPPRRRWRPFRPPLAAQIEDL
 RRSIVNLTFKQSPNPPPGPPPKKKKSAPKPKPTQPKKKKQQAARTKRKPKPGKRQRM
 CMKLES DKTFPIMLNGQVNGYACVVGGRLMKPLHVEGKIDNEQLAAVKLKKASMYDLE
 YGDVPQNMKSDTLQYTS DKPPGFYNWHHGAVQYENGRFTVPRGVGGKGDSGRPI LDNR
 GRVVAIVLGGANEGTRTALS VVTWNQKGV TIRDTPEGSEPWSLVTALCVLSNVTFPCD
 KPPVCYSLTPERTLDVLEENV DNP NYDTLLENVLKCPSRRPKRSITDDFTLTSPYLG F
 CPYCRHSTPCFSPIKIENVWDESDDGSIRIQVSAQFGYNQAGTADVTKFRYMSFDHHDH
 DIKEDSMEKIAISTSGPCRRLGHKGYFLLAQCPPGDSVTVSITSGASENSCTVEKKIR
 RKFVGREEYLFPPVHGKLVKCHVYDHLKETSAGYITMHRPGPHAYKSYLEEASGEVYI
 KPPSGKNVTYECKG DYSTGIVSTRTKMNGCTKAKQCIAYKSDQTKWVFNSPDLIRHT
 DHSVQGKLHIPFRLTPTVCPVPLAHTPTVTKWFKGITLHLTAMRPTLLTTRKLGLRAD
 ATAEWITGSTS RNFSV GREGLEYVWGNHEPVRVWAQESAPGDPHGWPHEII IHYYHRH
 PVYTVIVLCGVALAILVGTASSAACIAKARRDCLTPYALAPNATVPTALAVLCCIRPT
 NAETFGETLNHLWFNNQPFLWAQLCIPLAALVILFRCFSCCMPFLLVAGVCLGKVDAF
 EHATTVPNVPGIPYKALVERAGYAPLNLEITVVSSELTPSTNKEYVTCKFHTVIPSPQ

VKCCGSLECKASSKADYTCRVFGGVYPFMWGGAQCFCDSSENTQLSEAYVEFAPDCTID
 HAVALKVHTAALKVGLRIVYGN'TTAHLDTFVNGVTPGSSRDLKVIAGPISAAFSFPDH
 KVVIRKGLVYNYDFPEYGAMKPGAFGDIQASSLDATDIVARTDIRLLKPSVKNIHVPY
 TQAVSGYEMWKNNNGRPLQETAPFGCKIEVEPLRASNCAYGHIPI SIDIPDAAFVRSS
 ESPTILEVSTVADCIYSADFGGSLTLQYKADREGHCPVHSHSTTAVLKEATTHVTAV
 GSITLHFSTSSPQANFIVSLCGKKSTCNAECKPPADHII GEPHKVDQEFQAAVSKTSW
 NWLLALFGGASSLIVVGLIVLVCSMLINTR"

3'UTR

3870-4150

BASE	COUNT	ORIGIN
1	ggccctctag	agctcact
61	agcgttaaga	acttcaagag
121	taggtgacgt	agtagacacg
181	actttccacc	agttttaccct
241	gccgctggag	gccgttttcgg
301	tagtcaactt	gacttttcaaa
361	agaagaagag	tgctcctaag
421	agaggacgaa	acgcaagcct
481	cggacaagac	attttccgac
541	gaggaaggct	gatgaaacca
601	ccgtgaaatt	gaagaaggct
661	tgaaatcaga	cacgctgcag
721	acggcgcagt	ccagtatgag
781	gcgacagcgg	aagaccgatc
841	gtgcaaatga	gggcacgcgt
901	ccattaggga	taccccccga
961	tttcgaatgt	cacgtttccca
1021	gaacactcga	cgtgctcga
1081	acgtcttgaa	atgtccatca
1141	ccagtcccta	cctgggggttc
1201	taaaaattga	gaacgtgtgg
1261	cacaattcgg	ctacaatcag
1321	tcgaccacga	ccatgacatc
1381	gaccctgccg	tcgtcttggc
1441	acagtgtaac	cgtcagtatc
1501	agatcaggag	gaagtgtgtc
1561	tggtaaaagt	ccacgttttac
1621	acaggccagg	cccacacgcg
1681	ttaaaccacc	ttctggcaag
1741	gtatcgtgag	cacgcgaacg
1801	acaagagcga	ccaaacgaaa
1861	actcagtgca	aggtaaattg
1921	cgttagctca	cacgcctaca

```

1981 caatgcgacc aacattgctg acaacgagaa aattggggct gcgagcagac gcaacagcag
2041 aatggattac aggggtctaca tccaggaatt tttctgtggg gcgagaaggg ctggagtacg
2101 tatggggtaa ccatgaacca gtcagagtct gggcccagga gtcggcacca ggcgacccac
2161 atggatggcc gcatgagatc atcatccact attatcatcg gcatccagtc tacactgtca
2221 ttgtgctgtg tgggtgtcgt cttgctatcc tggtaggcac tgcattcatca gcagcttgca
2281 tcgccaaagc aagaagagac tgctgacgc catacgcgct tgcaccgaac gcaacgggtac
2341 ccacagcatt agcgggtttt tgctgcatte ggccaaccaa cgctgaaaca tttggagaaa
2401 ctttgaacca tctgtggttt aacaaccaac cgtttctctg ggcacagtgtg tgcattcctc
2461 tggcagcgct tgttattctg ttccgctgct tttcatgctg catgcctttt ttattggttg
2521 caggcgctctg cctggggaag gtagacgcct tcgaacatgc gacctgtgtg ccaaagtgtc
2581 cggggatccc gtataaggcg ttggtcgaac gcgcagggtta cgcgccactt aacctggaga
2641 tcacggtcgt ctcatcggaa ttaacacctt caactaacia ggagtacgtg acctgcaaat
2701 tccacacagt cattccttca ccacaagtta aatgctgcgg gtccctcgag tgcaaggcat
2761 cctcaaaggc ggattacaca tgccgcgttt ttggcgggtg gtacctttt atgtggggag
2821 gcgcacaatg cttctgtgac agtgagaaca cacaactgag tgaggcgtag gtcgagttcg
2881 ctccagactg cactatagat cagcgagtcg cactaaaagt tcacacagct gctctgaaag
2941 tcggcctgcg tatagtatac ggcaacacca ccgcgcacct ggatacgttt gtcaatggcg
3001 tcacgccagg ttcctcacgg gacctgaagg tcatagcagg gccgatatca gccgcttttt
3061 caccctttga ccataagggtc gtcacagaa aggggcttgt ttacaactac gacttccctg
3121 agtatggagc tatgaaacca ggagcggttc gcgatattca agcatcctcg cttgatgcta
3181 cagacatagt agcccgact gacatacggc tgctgaagcc ttctgtcaag aacatccacg
3241 tcccctacac ccaagcagta tcaggggatg aaatgtggaa gaacaactca ggacgacccc
3301 tgcaagaaac agcaccattt ggatgtaaaa ttgaagtgga gcctctgcca gcgtctaact
3361 gtgcttacgg gcacatccct atctcgattg acatccctga tgcagctttt gtgagatcat
3421 cagaatcacc aacaatttta gaagttagct gcacagtagc agactgcatt tattctgcag
3481 actttggtgg ttctctaaca ttacagtaca aagctgacag ggagggacat tgtccagttc
3541 actcccactc cagcagagct gttttgaagg aagcgaccac acatgtgact gccgtaggca
3601 gcataaact acatttttagc acatcgagcc cacaagcaaa ttttatagtt tcgctatgcg
3661 gcaagaagtc cacctgcaat gctgaatgta aaccaccggc cgaccacata attggagaac
3721 cacataaagt cgaccaagaa ttccaggcgg cagtttccaa aacatcttgg aactggctgc
3781 ttgcaactgt tgggggagca tcatccctca ttgttgtagg acttatagtg ttggtctgca
3841 gctctatgct tataaacaca cgtagatgac tgagcgcgga cactgacata gcggtaaaac
3901 tcgatgtact tccgaggaag cgtggtgcat aatgccacgc gccgcttgac actaaaactc
3961 gatgtatttc cgaggaagca cagtgcataa tgctgtgcag tgtcacatta atcgtatatc
4021 acactacata ttaacaacac tatatcactt ttatgagact cactatgggt ctctaataa
4081 cactacacat attttactta aaaacactat acacacttta taaattctct cataatttca
4141 ctttaggttt
//

```

[illegible]

1	cattgacgtc	aatgggtgga	gtatttacgg	taaactgcc	acttggcagt	acatcaagtg
61	tatcatatgc	caagtacgcc	ccctattgac	gtcaatgacg	gtaaatggcc	cgcttggcat
121	tatgcccagt	acatgacctt	atgggacctt	cctacttggc	agtacatcta	cgtatttagtc
181	atcgctatta	ccatggtgat	gcggttttgg	cagtacatca	atgggcgtgg	atagcggttt
241	gactcacggg	gatttccaag	tctccacccc	attgacgtca	atgggagttt	gttttggcac
301	caaaatcaac	gggactttcc	aaaatgtcgt	aacaactccg	ccccattgac	gcaaattgggc
361	ggtaggcgtg	tacggtggga	ggtctatata	agcagagctc	tctggctaac	tagagaacc
421	actgcttact	ggcttatcga	aattaatacg	actcactata	gggagaccca	agctggctag
481	cgtttaaaact	taagcttggg	accgagctca	tactggcagg	cctgatcatc	acgtctctgt
541	ccacgttagc	cgaaagcgtt	aagaacttca	agagcataag	agggagccca	atcacccctct
601	acggctgacc	taaatagggtg	acgtagtaga	cacgcacctt	cccaccgcc	gaatgtttcc
661	ataccctcag	ctgaactttc	caccagttta	ccctacaaat	ccgatggctt	accgagatcc
721	aaaccctcct	aggcgccgct	ggaggccggt	tgggcccccg	ctggctgctc	aaatcgaaga
781	tcttaggagg	tcgatagtca	acttgacttt	caaacaacga	tcacctaatc	cgccgccagg
841	tccaccgcc	aagaagaaga	agagtgtctc	taagccaaaa	cctactcagc	ctaaaaagaa
901	gaagcagcaa	gccaaagagga	cgaaacgcaa	gcctaaacca	gggaaacgac	aacgtatgtg
961	tatgaagtgtg	gagtcggaca	agacatttcc	gatcatgctg	aacggccaag	tgaatggata
1021	tgcttgcgtt	gtcggaggaa	ggtgtatgaa	accactccac	gttgaaggaa	aaattgataa
1081	tgaacattta	gcggccgtga	aactgaagaa	ggctagcatg	tacgacttgg	agtacggcga
1141	cgttccccag	aacatgaaat	cagacacgct	gcagtacacc	agcgacaaac	caccgggctt
1201	ctacaactgg	caccacggcg	cagtcagta	tgagaatggg	agatttaccg	taccgaagag

4177 gaatgtaaacc caccggccga ccacataatt ggagaaccac ataaagtctga ccaagaattc
4237 caggcggcag tttccaaaac atcttggaac tggctgcttg cactgtttgg gggagcatca
4297 tccctcattg ttgtaggact tatagtgttg gtctgcagct ctatgcttat aaacacacgt
4357 agatgactga gcgcggacac tgacatagcg gtaaaactcg atgtacttcc gaggaagcgt
4417 ggtgcataat gccacgcgaa gcttgacact aaaactcgat gtattttccga ggaagcacag
4477 tgcataatgc tgtgcagtgt cacattaatc gtatatcaca ctacatatta acaacactat
4537 atcacttttta tgagactcac tatgggtctc taatatacac tacacatatt ttactttaaaa
4597 acactataca cactttataa attcttttat aatttttctt ttgcttttaga gcacactggc

4657 ggccggttact agtggatccg agctctagag ggcc

4177-4643 pcDW-HX45

4644-4790 vector sequence //

TOGETHER